Expansion of signal transduction pathways in fungi by whole-genome duplication

Luis M. Corrochano[1], Alan Kuo[2], Marina Marcet-Houben[3,4], Silvia Polaino[5], Asaf Salamov[2], José M. Villalobos-Escobedo[6], M. Isabel Álvarez[7], Javier Avalos[1], Diane Bauer[2], Ernesto P. Benito[7,8], Isabelle Benoit[9], Gertraud Burger[10], Lola P. Camino[1], David Cánovas[1], Enrique Cerdá-Olmedo[1], Jan-Fang Cheng[2], Angel Domínguez[7], Marek Eliáš[11], Arturo P. Eslava[7], Fabian Glaser[12], Jane Grimwood[2,13], Gabriel Gutiérrez[1], Joseph Heitman[14], Bernard Henrissat[15,16], Enrique A. Iturriaga[7], B. Franz Lang[10], José L. Lavín[17], Soo Chan Lee[14], Wenjun Li[14], Erika Lindquist[2], Sergio López-García[18], Eva M. Luque[1], Ana T. Marcos[1], Joel Martin[2], Kevin McCluskey[19], Humberto R. Medina[1], Alejandro Miralles-Durán[1], Atsushi Miyazaki[20], Elisa Muñoz-Torres[21], José A. Oguiza[22], Robin A. Ohm[2], María Olmedo[1], Margarita Orejas[23], Lucila Ortiz-Castellanos[24], Antonio G. Pisabarro[22], Julio Rodríguez-Romero[1], José Ruiz-Herrera[24], Rosa Ruiz-Vázquez[18], Catalina Sanz[7], Wendy Schackwitz[2], Jeremy Schmutz[2,13], Mahdi Shahriari[7], Ekaterina Shelest[25], Fátima Silva-Franco[18], Darren Soanes[26], Khajamohiddin Syed[27], Víctor G. Tagua[1], Nicholas J. Talbot[26], Michael R. Thon[7,8], Hope Tice[2], Ronald P. de Vries[9], Ad Wiebenga[9], Jagjit S. Yadav[27], Edward L. Braun[28], Scott Baker[29], Victoriano Garre[18], Benjamin A. Horwitz[30], Santiago Torres-Martínez[18], Alexander Idnurm[5,31], Alfredo Herrera-Estrella[6], Toni Gabaldón[3,4,32], Igor V. Grigoriev[2]

[1]Department of Genetics, University of Seville, 41012 Seville, Spain

[2]Department of Energy Joint Genome Institute, Walnut Creek, CA 94598, USA

[3]Centre for Genomic Regulation (CRG), 08003 Barcelona, Spain

[4]Universitat Pompeu Fabra (UPF), 08003 Barcelona, Spain

[5]School of Biological Sciences, University of Missouri-Kansas City, Kansas City, MO 64110,

USA

[6]Laboratorio Nacional de Genómica para la Biodiversidad (LANGEBIO), Centro de Investigación y Estudios Avanzados, 36821 Irapuato, Guanajuato, México

[7]Departamento de Microbiología y Genética, Universidad de Salamanca, 37007 Salamanca, Spain

[8]Instituto Hispano-Luso de Investigaciones Agrarias (CIALE), Universidad de Salamanca, 37185 Salamanca, Spain

[9]CBS-KNAW Fungal Biodiversity Centre & Fungal Molecular Physiology, Utrecht University, 3584 CT Utrecht, The Netherlands

[10]Universite de Montreal, Biochimie, Montreal, Québec H3C 3J7 Canada

[11]Department of Biology and Ecology, Faculty of Science, University of Ostrava, 710 00 Ostrava, Czech Republic

[12]Lokey Interdisciplinary Center for Life Sciences and Engineering, Technion - Israel Institute of Technology, Haifa 3200003, Israel

[13]HudsonAlpha Institute of Biotechnology, Huntsville, AL 35806, USA

[14]Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, NC 27710, USA

[15]Centre National de la Recherche Scientifique (CNRS), UMR7257, Université Aix-Marseille, 13288 Marseille, France

[16]Department of Biological Sciences, King Abdulaziz University, Jeddah 21589, Saudi Arabia

[17]Genome Analysis Platform, CIC bioGUNE, Bizkaia Technology Park, 48160 Derio, Bizkaia, Spain

[18]Departamento de Genética y Microbiología, Universidad de Murcia, 30071 Murcia, Spain

[19]Department of Plant Pathology, Kansas State University, Manhattan, KS 66506, USA

[20]Department of Biological Sciences, Faculty of Science and Engineering, Ishinomaki Senshu

University, Ishinomaki 986-8580, Japan

[21]Departamento de Biología Celular y Patología, Facultad de Medicina, Universidad de Salamanca, 37007 Salamanca, Spain

[22]Department of Agrarian Production, Public University of Navarre, 31006 Pamplona, Spain

[23]Instituto de Agroquímica y Tecnología de Alimentos, Consejo Superior de Investigaciones Científicas (IATA-CSIC), 46100 Burjassot, Valencia, Spain

[24]Departamento de Ingeniería Genética, Unidad Irapuato, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, 36821 Irapuato, Gto., Mexico

[25]Leibniz Institute for Natural Product Research and Infection Biology (Hans Knoell Institute), 07745 Jena, Germany

[26]Biosciences, College of Life & Environmental Sciences, University of Exeter, Exeter EX4 4QD, UK

[27]Department of Environmental Health, University of Cincinnati College of Medicine, Cincinnati, OH 45267-0056, USA

[28]Department of Biology, University of Florida, Gainesville, FL 32611, USA

[29]Pacific Northwest National Laboratory, Richland, WA 99352, USA

[30]Department of Biology, Technion - Israel Institute of Technology, Haifa 32000, Israel

[31]School of BioSciences, University of Melbourne, VIC 3010, Australia

[32]Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain.

Correspondence to: L.M. Corrochano. Email: corrochano@us.es.

Summary

Plants and fungi use light and other signals to regulate development, growth, and metabolism. The fruiting bodies of the fungus *Phycomyces blakesleeanus* are single cells that react to environmental cues, including light, but the underlying mechanisms are largely unknown [1]. The related fungus *Mucor circinelloides* is an opportunistic human pathogen that changes its mode of growth upon receipt of signals from the environment to facilitate pathogenesis [2]. Understanding how these organisms respond to environmental cues should provide new insights into the mechanisms of sensory perception and signal transduction by a single eukaryotic cell, and their role in pathogenesis. We sequenced the genomes of *P. blakesleeanus* and *M. circinelloides*, and show that they have been shaped by a whole genome duplication (WGD), which is rarely observed in fungi [3-6]. We show that the WGD has resulted in expansion of gene families, including those involved in signal transduction, and that duplicated genes have specialized following WGD, as shown by differences in their regulation by light. The transcriptional response to light indeed varies with the developmental stage and is still observed in a photoreceptor mutant of *P. blakesleeanus*. A phototropic mutant of *P. blakesleeanus* with a heterozygous mutation in the photoreceptor gene *madA* shows that photosensor dosage is important for the magnitude of signal transduction. We conclude that the WGD provided the means to improve signal transduction for enhanced perception of environmental signals. Our results will help to understand the role of genome dynamics in the evolution of sensory perception in eukaryotes.

Results and Discussion

Whole-genome Duplications in the Evolution of the Mucoromycotina Fungi

Gene duplication has expanded the number of genes for photoreception in *P. blakesleeanus* and *M. circinelloides* (Figure 1A, 1B) [7-9], and we hypothesize that gene duplications and specialization may have provided new proteins to expand the sensory repertoire of these fungi. The 53.9 Megabase (Mb) *P. blakesleeanus* and the 36.6 Mb *M. circinelloides* genomes, and their respective mitochondrial DNAs (mtDNAs), were sequenced using a Sanger whole-genome shotgun approach, and improved using targeted finishing (Supplemental Experimental Procedures; Table S1A-D; Figure S1). The increase in genome size in *P. blakesleeanus* is, in part, due to a large amount of repetitive DNA, including transcribed transposable elements (Table S1E-G). 16,528 (*P. blakesleeanus*) and 11,719 (*M. circinelloides*) protein-coding genes were annotated and compared to homologous proteins from other fungi (Figure 1C). Comparison of the two genomes with that of *R. delemar* and other fungi suggests that a whole-genome duplication (WGD) occurred early in the Mucoromycotina lineage. The fungal kingdom contains the subkingdom Dikarya and a number of early divergent lineages that include the Mucoromycotina with *P. blakesleeanus*, *M. circinelloides*, and *Rhizopus delemar* as prominent members [10]. Only two WGDs have been reported in fungi: in the Saccharomycotina, a member of the Dikarya [4-6] and in *R. delemar* [3].

Ancient WGDs are difficult to detect because gene loss and rearrangements have resulted in the absence of clear regions of synteny. However, analysis of genomes from Mucoromycotina species shows that they have more members per gene family than the genomes from Dikarya fungi (2.9-3.6 versus 1.6–2.2) (Supplemental Experimental Procedures; Table S2A). Mucoromycotina species, for example, have a large fraction of gene families with a number of members that is greater than the average in fungi (50%-68% versus

6.9%-22% for each Dikarya species) (Table S2A). Moreover, Mucoromycotina genomes have more duplicated regions than other fungal genomes (Table S2B). Duplicated regions in Mucoromycotina fungi correspond to 8-77% of their genomes and include 4-13 genes in average (Tables S2B and S2C). The presence of such amounts of duplicated DNA in Mucoromycotina genomes can be best explained by a WGD in the last common ancestor of Mucoromycotina fungi followed by a subsequent WGD in the *Rhizopus* lineage. The alternative explanation, the occurrence of lineage-specific gene duplications, is less likely since gene loss is known to be rampant. We confirmed the presence of large amounts of duplicated regions after WGDs in the genomes of *S. cerevisiae* and *R. delemar* as expected, but we noticed large amounts of duplicated DNA in the genomes of *Mortierella alpina*, *Cochliobolus heterostrophus*, *Puccinia graminis*, and *Laccaria bicolor* (Tables S2B and S2C). It has been shown that the *P. graminis*, and *L. bicolor* genomes have expanded lineage-specific gene families that are possibly involved in pathogenesis and symbiosis [11, 12]. These duplicated regions contain large fractions of lineage-specific genes (Table S2C) supporting the proposal that duplicated regions in non-Mucoromycotina genomes arose after species-specific segmental duplications, but only additional analysis will help to rule out the occurrence of further WGDs. Additional WGD signatures can be observed in families of three genes from genome pairs. These types of duplicates are more frequent in the genomes of Mucoromycotina species than in non-Mucoromycotina species, suggesting that the former harbors traces of past WGDs (Table S2D).

To gain further insight into past genome expansions in the Mucoromycotina, we reconstructed the complete collection of evolutionary histories (*i.e.* the phylome) for genes within Mucoromycotina fungi with 13 other sequenced fungal genomes (Supplemental Experimental Procedures). The gene trees were analyzed to detect and date duplication events [13] (Figure 1C). This method has been used successfully to detect and characterize the origin

and nature of the WGD that took place in the *S. cerevisiae* lineage [14]. In addition to the recent WGD described in *R. delemar* (0.43 duplications per gene), we detected a much larger duplication peak (0.70-0.96 duplications per gene) in the lineage preceding the Mucoromycotina species. This result is consistent with the occurrence of at least one earlier WGD preceding the diversification of the Mucoromycotina species analyzed. This early WGD explains gene duplications in the oxidative phosphorylation complex in Mucorales [15] and segmental duplications in *Lichtheimia corymbifera* where previously a species-specific WGD had been rejected [16].

Expansion of Gene Families after WGD.

A consequence of WGD is expansion of certain gene families. Duplicated genes in the four Mucoromycotina species, for example, contained an abundance of the Gene Ontology (GO) terms: protein kinase activities (GO:4674, GO:4672, GO:4713, GO:4707; P-value at least $1x10^{-6}$), fructose 2,6-bisphosphate metabolic process (GO:6003; P=1.3 $x10^{-6}$), ATP-binding (GO:5524; P=1.8 $x10^{-46}$), and protein transport (GO:15031; P=1.6 $x10^{-17}$), suggesting duplications of genes encoding signalling pathways and transport components (Supplemental Experimental Procedures; Figure 2; Table S3). This is supported by further analysis of the abundance of signal transduction genes. We have limited this analysis to the genomes of *P. blakesleeanus*, *M. circinelloides*, and *R. delemar* as they were the only Mucoromycotina genomes available when we started the project.

A versatile and widespread mode of signaling starts at the cell surface with seven-transmembrane (7TM) receptors that are coupled to heterotrimeric G proteins. The G protein trimer is composed of a Gα subunit and a Gβγ heterodimer that dissociates from Gα upon activation. Heterotrimeric G-protein signalling is central to the life cycle and virulence of fungi [17, 18]. All the gene families encoding the subunits of heterotrimeric G proteins are

expanded. For example, the Gα subunit family contains 10, 12 and 12 genes in *P. blakesleeanus*, *M. circinelloides*, and *R. delemar*, respectively, compared to an average of three in the Ascomycota or seven in the Dikarya (Figure 2A). Like the Gα subunits, Gβ and Gγ subunit families are expanded in all three species (Figure 2A). A single Gβ subunit gene has been found in the Ascomycota as well as in the basidiomycetes *U. maydis* and *C. neoformans* [19, 20]. In *P. blakesleeanus* we identified five Gβ genes, in *M. circinelloides* three, and in *R. delemar* four (Figure 2A). Theoretically, a very large number of G protein heterotrimers could be built from the 10-12, 5-3 and 4-3 α, β and γ subunits in *P. blakesleeanus* and *M. circinelloides*, respectively. An estimate of 21, 21 and 10 genes for G protein coupled receptors (GPCR) in *P. blakesleeanus*, *M. circinelloides* and *R. delemar*, respectively (Table S3A), suggests moderate expansion as 10-12 genes for GPCRs are known in Dikarya [21].

The number of other signal transduction genes has increased compared to the Dikarya, including genes for protein kinases, TRAFAC class GTPases, and regulators of GTPases of the Ras superfamily (Supplemental Experimental Procedures; Tables S3A-D). For example, the expansion of kinase families relative to Dikarya is 3-4 fold, as there are 63, 70 and 84 CAMK kinase genes in *P. blakesleeanus*, *M. circinelloides*, and *R. delemar* respectively, compared to 22 and 21 in *Neurospora crassa* and *Ustilago maydis*, respectively. Other kinase families show larger expansion, *e.g.* there are 11-18 genes for casein kinase 1 in Mucoromycotina fungi compared to 2-3 in Dikarya (Figure 2A; Table S3C).

Some but not all families in a given category are expanded, *e.g.* for photoreception [22], the genes for components of the WC photoreceptor complex (WC-1 and WC-2) are duplicated, but not the cryptochrome gene; genes for casein kinase 1 are duplicated, but not those encoding the sensor histidine kinases (Figure 2A). Genes for calcium sensory or pH sensory pathways show non-uniform duplication: in the calcium pathway there are multiple

calmodulin genes and three calcineurin catalytic subunit genes, but only a single calcineurin regulatory subunit; in the pH pathway there are 3-4 genes for the PacC transcription factor and two genes for PalA, but one gene for PalB or PalC as in the Ascomycete *Aspergillus nidulans* [23] (Table S3A). There are expanded cyclin families, but only a single identified mitotic cyclin, compared to several in the Dikarya (Figure 2A). Not all gene families have expanded. For example, the genes for proteins that participate in genome defence through RNAi have not duplicated after WGD (Figure 2A). Their presence suggests the existence of RNAi-mediated regulation in *P. blakesleeanus* and *R. delemar*, as described in *M. circinelloides* [24].

The WGD have multiplied genes involved in cell wall biosynthesis, in particular genes for chitin synthases and chitin deacetylases (Figure 2A). These enzymes may have specialized to modulate the growth response of the sporangiophore after environmental stimuli. Functional specialization after gene duplication should have played a key role in *M. circinelloides* and other pathogenic Mucoromycotina fungi. A reduction in pathogenesis is observed in mutants of the photoreceptor gene *wc-1* of *Fusarium oxysporum* and *Cryptococcus neoformans* [25, 26]. Since *M. circinelloides wc* genes have specialized their sensory role after gene duplication it is possible that some of the WC proteins now serve as pathogenicity factors as suggested by preliminary experiments (Torres-Martínez, pers. comm.). Calcineurin is a virulence factor in several fungi, including *M. circinelloides* [27]. The number of calcineurin A catalytic subunit genes have increased with three genes in *M. circinelloides* compared to one in other fungi (Table S3A). One of them, *cnaA*, is involved in virulence [27], further confirming gene specialization to facilitate pathogenesis after duplication.

The expansion of gene families that we have observed in Mucoromycotina fungi can be accounted for by a WGD and retention of the resulting paralogs. The expansion of some families, for example chitin deacetylases or Gα subunits (Figure 2A, Table S3A), cannot be

explained by a WGD alone, and additional segmental duplications appear to have occurred. The increase in the number of proteins that participate in signal transduction pathways and cell wall biosynthesis after WGD should have helped Mucoromycotina fungi to improve environmental sensing and responses, including the perception of potential hosts for pathogenic fungi. The biological role of the expanded gene in set will require further genetical and biochemical characterization.

Duplicated Genes Differ in their Transcriptional Response to Light

In order to investigate whether duplicated genes have specialized after WGD we asked if duplicated genes from *P. blakesleeanus* and *M. circinelloides* responded differently to environmental cues, like light. We characterized the transcriptome in cultures kept in the dark or after exposure to 30 min of blue light ($2.3/10^3$ J/m^2) and we observed that duplicated genes have specialized their pattern of expression following WGD and subsequent divergence (Supplemental Experimental Procedures; Figure 2B). Most of the genes encoding components of the photoreceptor complex (WCC) or the regulatory subunits of protein kinase A showed a similar expression pattern in *P. blakesleeanus* and *M. circinelloides* with only some genes regulated by light despite being transcriptionally active in both mycelia and sporangiophores (Figure 2B). In addition, we noted that genes encoding proteins of relevance in signal transduction, such as the photoreceptor WcoB, the G protein beta subunits Gpb1 and Gpb3, and the kinases Pkac-1, Mps1-1, CK1-7 and CK1-8, showed opposite responses to light in *P. blakesleeanus* and *M. circinelloides*. Our results confirm that duplicated genes have different patterns of expression in different organisms, as well as between specific tissues in the same organism after WGD.

A Refined Transcriptional Response to Light in *P. blakesleeanus*

The expansion in the number of photoreceptor genes and other signal transduction proteins may have allowed the fine-tuning of the response to light, for example allowing tissue-specific transcriptional responses to light. We thus characterized in more detail the global transcriptional response to light in *P. blakesleeanus* mycelium and sporangiophores by RNAseq (Supplemental Experimental Procedures; Figure 3; Table S4). A total of 2,024 genes were responsive to light in the mycelium (1,421 induced and 603 repressed, about 12% of the protein-coding genes), compared to 1,212 genes in the sporangiophores (1,042 induced and 170 repressed). The transcriptional response to light was specific for each developmental stage because only 113 genes were regulated by light in both the mycelium and sporangiophores (Figure 3A-B). We conducted the same analysis on a *madA madB* double mutant (strain L51) that is considered blind [28]. Only 159 genes were regulated by light in the mycelium of the *madA madB* mutant, which confirmed the relevance of the Mad complex in light-dependent gene regulation (Figure 3A). Surprisingly, the *madA madB* mutant showed a significant response to light in sporangiophores since 3,513 genes were regulated by light, consistent with the activity of other photoreceptors (Figure 3A). It is noteworthy that most responsive genes in the *madA madB* mutant sporangiophores were those repressed by light (Figure 3A). This suggests the activity of light-dependent repressors in the absence of the Mad complex as proposed by EMSA experiments that showed the binding of proteins in the dark to a light-regulated promoter in a *madA madB* mutant [29]. Seven gene clusters were enriched in regulatory genes in light-induced mRNAs from the mycelium of the wild type, and in ribosome biogenesis genes in the sporangiophores (Figure 3C). The *P. blakesleeanus* and *M. circinelloides* genomes contain 879 and 650 genes for transcription factors (TFs), respectively (about 5% of the protein-coding genes), with an abundance of C_2H_2 Zn finger TFs (Fig. S2). In fact, light regulates 9% of the *P. blakesleeanus* transcription factor genes (Table S3E-G and S4E-F), and we propose that the stage-specific transcriptional response to

light relies on the expanded set of photoreceptors and other light-dependent transcriptional regulators that arose after WGDs. Specialization of genes for signal transduction following WGD has been observed in vertebrate vision [30]. Strikingly, in both vertebrates and fungi the expansion of signal transduction genes after WGD has resulted in improved sensory perception.

Reduced Sensitivity to Light in Strains with Wild Type and *madA* Mutant Nuclei

The photoresponse in the *P. blakesleeanus madA madB* mutant suggested the action of additional photoreceptors. We therefore searched for the presence of potential candidates. The perception of light in *madI* mutants is reduced 10 to 1000-fold, and this gene may therefore encode a secondary photoreceptor [31, 32] (Figure 4A). To identify *madI* we crossed two *madI* strains with a wild-type strain, and characterized the phototropism and linked PCR-RFLP markers in the progeny [33]. Weak linkage was found for the *madI* mutation and three scaffolds, including the one carrying *madA* (Supplemental Experimental Procedures; Table S5). We therefore sequenced the genomes of two *madI* mutants (L151 and L153), along with another 17 *mad* mutant strains, and the sequences were scanned across the three scaffolds. We found that the two *madI* strains had an identical and unique mutation in *madA* [8] changing a conserved proline to leucine (Figure 4B-C). However, the *madI* strains also contained the wild-type allele, indicating that they were heterozygous for this gene. To confirm these observations we sequenced *madA* in 63 *madI* x wild type progeny: eight only had the mutation in the *madA* gene, 35 were wild type, and 20 were heterozygotes. The two *madI* mutants had a high number of heterozygous sites across their genomes compared to other strains as shown by the ratio of heterozygous to non-heterozygous single nucleotide polymorphisms (SNPs) (Figure 4D). Analysis of all scaffolds in the L151 and L153 genomes showed that the heterozygous SNPs are distributed throughout all chromosomes (Figure S3),

suggesting that the two *madI* strains are heterokaryons or diploids, rather than being aneuploid or carrying a segmental duplication. The observation that two of the *madI* strains are heterozygous wild type/*madA* mutants shows that sensitivity to light is related to the dosage of the MadA photoreceptor.

When considered together, our characterization of the genomes of *P. blakesleeanus* and *M. circinelloides*, our comparative fungal genome analysis, and our subsequent gene functional studies provide new insight into the occurrence and consequences of WGDs in the evolution of fungi. Expansion and specialization of genes for signal transduction and cell-wall biosynthesis following WGDs in the Mucoromycotina has provided new proteins that have enabled these fungi to refine the way they perceive signals from the environment to regulate their growth and development. Our results provide new genomic tools to unravel the molecular mechanisms of sensory perception in early diverging fungi that will help to understand the evolution of sensory perception in eukaryotes.

Author contributions

A.K., M.M.-H., S.P. and A.S. contributed equally to this work. L.M.C. coordinated the project. E.L.B., S.B., S.T.-M. and L.M.C. conceived the project. I.V.G. directed all genome-sequencing and analysis efforts at JGI. J.R.-R., V.G.T., A.M.-D., C.S., M.S., E.M.L., F.S.-F. and V.G. provided DNA and RNA samples. A.K., A.S., E.L., J.-F.C., W.S., J.M., J.S., J.G., H.T. and D.B. sequenced, assembled and annotated the genomes. M.M.-H, A.S., T.G. and I.V.G. did genome comparisons and phylogenomics analysis. M.I.A., J.A., E.P.B., I.B., G.B., L.P.C., D.C., E.C.O., L.M.C., A.D., M.E., A.P.E., V.G., F.G., G.G., J.H., B.H, B.A.H., A.I., E.A.I., B.F.L., J.L.L., S.C.L., W.L., A.T.M., H.R.M., A.M., E.M.-T., J.A.O., R.A.O., M. Olmedo, M. Orejas, L.O.-C., A.G.P., J.R-R., J.R.-H., R.R.-V., C.S., M.S., E.S., F.S.-F., D.S., K.S., V.G.T., N.J.T., M.R.T., S.T.-M., R.P.V., A.W. and J.S.Y. analyzed the genomes.

J.M.V.-E., E.M.L., S.L.-G., F.S.-F., R.R.-V., S.T.-M., V.G., L.M.C. and A.H.-E performed the transcriptomic analysis. S.P., S.B., K.M. and A.I. performed the genetic characterization of *madI*. L.M.C., B.A.H., S.T.-M, A.I., A.H.-E., T.G. and I.V.G. wrote the paper with input from all authors.

Data deposition

Genome assemblies and annotations have been deposited in DDBJ/EMBL/GenBank under the accessions AMYC00000000 (*P. blakesleeanus*) and AMYB00000000 (*M. circinelloides*), and are accessible via the JGI web page (http://jgi.doe.gov). The mtDNAs have been deposited in DDBJ/EMBL/GenBank under the accessions KR809878 (*P. blakesleeanus*) and KR809877 (*M. circinelloides*). The phylomes can be accessed at phylomeDB (http://phylomedb.org) with phylomeIDs 252, 253, 254, 255 and 256. The gene expression results have been deposited in the GEO database with accessions GSE64369 (*P. blakesleeanus*) and GSE58264 (*M. circinelloides*).

Acknowledgments

This article is dedicated to the memory of Max Delbrück (1906-1981) for promoting and supporting the use of *Phycomyces* as a model organism for research on the mechanisms of signal transduction. The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. The annotation and coordination of the project was supported by European funds (European Regional Development Fund, ERDF), the Spanish Ministry of Education and Science (BIO2005-25029-E), and the Regional Government (Junta de Andalucía, P06-CVI-01650) to LMC. The RNAseq project was supported by Conacyt (Mexico) with grant FORDECYT-2012-02-193512 to AHE. Genetic analysis was supported

by the US National Science Foundation grant MCB-0920581 to AI. We thank Andrii Gryganskyi, Daniele Armaleo, and Paul Dyer for allowing us to use the genome sequences from *Lichtheimia corymbifera*, *Cladonia grayi*, and *Xanthoria parietina*, respectively, prior to publication. We thank M. M. Gil Sánchez for *P. blakesleeanus* photographs in Fig. 1A.

References

1. Cerdá-Olmedo, E. (2001). *Phycomyces* and the biology of light and color. FEMS microbiology reviews *25*, 503-512.
2. Corrochano, L.M., and Garre, V. (2010). Photobiology in the Zygomycota: multiple photoreceptor genes for complex responses to light. Fungal Genet Biol *47*, 893-899.
3. Ma, L.J., Ibrahim, A.S., Skory, C., Grabherr, M.G., Burger, G., Butler, M., Elias, M., Idnurm, A., Lang, B.F., Sone, T., et al. (2009). Genomic analysis of the basal lineage fungus *Rhizopus oryzae* reveals a whole-genome duplication. PLoS Genet *5*, e1000549.
4. Wolfe, K.H., and Shields, D.C. (1997). Molecular evidence for an ancient duplication of the entire yeast genome. Nature *387*, 708-713.
5. Dietrich, F.S., Voegeli, S., Brachat, S., Lerch, A., Gates, K., Steiner, S., Mohr, C., Pöhlmann, R., Luedi, P., Choi, S., et al. (2004). The *Ashbya gossypii* genome as a tool for mapping the ancient *Saccharomyces cerevisiae* genome. Science *304*, 304-307.
6. Kellis, M., Birren, B.W., and Lander, E.S. (2004). Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. Nature *428*, 617-624.
7. Sanz, C., Rodríguez-Romero, J., Idnurm, A., Christie, J.M., Heitman, J., Corrochano, L.M., and Eslava, A.P. (2009). *Phycomyces* MADB interacts with MADA to form the

primary photoreceptor complex for fungal phototropism. Proc Natl Acad Sci USA *106*, 7095-7100.

8. Idnurm, A., Rodríguez-Romero, J., Corrochano, L.M., Sanz, C., Iturriaga, E.A., Eslava, A.P., and Heitman, J. (2006). The *Phycomyces madA* gene encodes a blue-light photoreceptor for phototropism and other light responses. Proc Natl Acad Sci USA *103*, 4546-4551.

9. Silva, F., Torres-Martínez, S., and Garre, V. (2006). Distinct *white collar-1* genes control specific light responses in *Mucor circinelloides*. Mol Microbiol *61*, 1023-1037.

10. McLaughlin, D.J., Hibbett, D.S., Lutzoni, F., Spatafora, J.W., and Vilgalys, R. (2009). The search for the fungal tree of life. Trends Microbiol *17*, 488-497.

11. Martin, F., Aerts, A., Ahren, D., Brun, A., Danchin, E.G., Duchaussoy, F., Gibon, J., Kohler, A., Lindquist, E., Pereda, V., et al. (2008). The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. Nature *452*, 88-92.

12. Duplessis, S., Cuomo, C.A., Lin, Y.C., Aerts, A., Tisserant, E., Veneault-Fourrey, C., Joly, D.L., Hacquard, S., Amselem, J., Cantarel, B.L., et al. (2011). Obligate biotrophy features unraveled by the genomic analysis of rust fungi. Proc Natl Acad Sci U S A *108*, 9166-9171.

13. Huerta-Cepas, J., and Gabaldón, T. (2011). Assigning duplication events to relative temporal scales in genome-wide studies. Bioinformatics *27*, 38-45.

14. Marcet-Houben, M., and Gabaldon, T. (2015). Beyond the Whole-Genome Duplication: Phylogenetic Evidence for an Ancient Interspecies Hybridization in the Baker's Yeast Lineage. PLoS Biol *13*, e1002220.

15. Marcet-Houben, M., Marceddu, G., and Gabaldón, T. (2009). Phylogenomics of the oxidative phosphorylation in fungi reveals extensive gene duplication followed by functional divergence. BMC evolutionary biology *9*, 295.

16. Schwartze, V.U., Winter, S., Shelest, E., Marcet-Houben, M., Horn, F., Wehner, S., Linde, J., Valiante, V., Sammeth, M., Riege, K., et al. (2014). Gene expansion shapes genome architecture in the human pathogen *Lichtheimia corymbifera*: an evolutionary genomics analysis in the ancient terrestrial mucorales (Mucoromycotina). PLoS Genet *10*, e1004496.

17. Lafon, A., Han, K.H., Seo, J.A., Yu, J.H., and d'Enfert, C. (2006). G-protein and cAMP-mediated signaling in aspergilli: a genomic perspective. Fungal Genet Biol *43*, 490-502.

18. Li, L., Wright, S.J., Krystofova, S., Park, G., and Borkovich, K.A. (2007). Heterotrimeric G protein signaling in filamentous fungi. Annu Rev Microbiol *61*, 423-452.

19. Müller, P., Leibbrandt, A., Teunissen, H., Cubasch, S., Aichinger, C., and Kahmann, R. (2004). The Gbeta-subunit-encoding gene bpp1 controls cyclic-AMP signaling in *Ustilago maydis*. Eukaryot Cell *3*, 806-814.

20. Wang, P., Perfect, J.R., and Heitman, J. (2000). The G-protein beta subunit GPB1 is required for mating and haploid fruiting in *Cryptococcus neoformans*. Mol Cell Biol *20*, 352-362.

21. Xue, C., Hsueh, Y.P., and Heitman, J. (2008). Magnificent seven: roles of G protein-coupled receptors in extracellular sensing in fungi. FEMS microbiology reviews *32*, 1010-1032.

22. Idnurm, A., Verma, S., and Corrochano, L.M. (2010). A glimpse into the basis of vision in the kingdom Mycota. Fungal Genet Biol *47*, 881-892.

23. Peñalva, M.A., Tilburn, J., Bignell, E., and Arst, H.N., Jr. (2008). Ambient pH gene regulation in fungi: making connections. Trends Microbiol *16*, 291-300.

24. Nicolás, F.E., Moxon, S., de Haro, J.P., Calo, S., Grigoriev, I.V., Torres-Martínez, S., Moulton, V., Ruiz-Vázquez, R.M., and Dalmay, T. (2010). Endogenous short RNAs generated by Dicer 2 and RNA-dependent RNA polymerase 1 regulate mRNAs in the basal fungus *Mucor circinelloides*. Nucleic Acids Res *38*, 5535-5541.

25. Ruiz-Roldán, M.C., Garre, V., Guarro, J., Mariné, M., and Roncero, M.I. (2008). Role of the white collar 1 photoreceptor in carotenogenesis, UV resistance, hydrophobicity, and virulence of *Fusarium oxysporum*. Eukaryot Cell *7*, 1227-1230.

26. Idnurm, A., and Heitman, J. (2005). Light controls growth and development via a conserved pathway in the fungal kingdom. PLoS Biol *3*, e95.

27. Lee, S.C., Li, A., Calo, S., and Heitman, J. (2013). Calcineurin plays key roles in the dimorphic transition and virulence of the human pathogenic zygomycete *Mucor circinelloides*. PLoS pathogens *9*, e1003625.

28. Lipson, E.D., and Terasaka, D.T. (1981). Photogeotropism in *Phycomyces* double mutants. Experimental mycology *5*, 101-111.

29. Sanz, C., Benito, E.P., Orejas, M., Alvarez, M.I., and Eslava, A.P. (2010). Protein-DNA interactions in the promoter region of the *Phycomyces carB* and *carRA* genes correlate with the kinetics of their mRNA accumulation in response to light. Fungal Genet Biol *47*, 773-781.

30. Larhammar, D., Nordström, K., and Larsson, T.A. (2009). Evolution of vertebrate rod and cone phototransduction genes. Philos Trans R Soc Lond B Biol Sci *364*, 2867-2880.

31. Campuzano, V., Díaz-Mínguez, J.M., Eslava, A.P., and Alvarez, M.I. (1990). A new gene (*madI*) involved in the phototropic response of *Phycomyces*. Mol Gen Genet *223*, 148-151.

32. Alvarez, M.I., Eslava, A.P., and Lipson, E.D. (1989). Phototropism mutants of *Phycomyces blakesleeanus* isolated at low light intensity. Experimental mycology *13*, 38-48.

33. Chaudhary, S., Polaino, S., Shakya, V.P., and Idnurm, A. (2013). A new genetic linkage map of the zygomycete fungus *Phycomyces blakesleeanus*. PLoS One *8*, e58931.

Figure legends

Figure 1. Sensory perception and a whole genome duplication in the Mucoromycotina.
(A) The fruiting bodies, sporangiophores, of *Phycomyces blakesleeanus* grow out of the mycelium and reach several cm in length. The speed and direction of growth is controlled by signals from the environment including light, gravity, touch, wind, and the presence of nearby objects. The ball at the top of each fruiting body is the sporangium with spores. The direction of light is indicated by an arrow.
(B) The sporangiophores of *Mucor circinelloides* are small (about 1 mm) and show phototropism. The direction of light is indicated by an arrow.
(C) Evidence for a WGD in the Mucoromycotina. A fungal evolutionary tree with bootstrap support values lower than 95% indicated in black numbers at the branches. The average duplication per gene in each lineage is shown with a colour that indicates the phylome used for the duplication density calculation. The branch where the WGD took place is marked in red with a dot and an arrow. The graph (scale on the bottom) represents the percentage of genes in a given species that belong to one of the following categories: yellow, protein present in all species; light yellow, ancestral proteins that have homologs in the outgroups; brown, Fungi-specific proteins; green, Mucoromycotina-specific proteins which appear in all four species; light green, Mucoromycotina-specific proteins; grey, species-specific proteins. The red bars (scale on top) represent the total number of proteins encoded in each genome.

Figure 2. Gene expansion and transcriptional specialization in the Mucoromycotina.
(A) Gene abundance in Mucoromycotina, *Neurospora crassa* and Dikarya fungi. The x-axis indicates number of genes, and the bars, from bottom to top, indicate numbers of predicted genes for the three Mucoromycotina (Pb, *P. blakesleeanus*; Mc, *M. circinelloides*; Rd, *R.*

delemar), and average number of genes for *N. crassa* (Nc), and Dikarya (Table S3). The three *wc-1* genes are *madA*, *wcoA*, and *wcoB*, and the four *wc-2* genes are *madB*, *wctB*, *wctC*, and *wctD*.

(B) Expression patterns in response to light of expanded genes in *P. blakesleeanus* and *M. circinelloides*. Differential expression of genes was obtained for two *P. blakesleeanus* stages (mycelium and sporangiophore), and *M. circinelloides* mycelium. Results are represented with the logarithm base ten of FDR, (FDR<0.05; Fold-change>2). M= Mycelium, S= Sporangiophore, L= Light, D= Dark (Table S4).

Figure 3. The influence of light on gene expression in two developmental stages of *P. blakesleeanus.*

(A) Differential expression (light/dark) in the mycelium or the sporangiophore of the wild-type and the *madA madB* mutant strain (L51). Differentially expressed genes with FDR ≤ 0.05 are shown in red. (Table S4).

(B) Overlap of genes induced and repressed in the wild type and mutant using RNA from mycelia (WT-ML and L51-ML) or sporangiophores (WT-SL and L51-SL).

(C) Category enrichment in differentially expressed genes (*FDR<0.05; **FDR<0.01). Each vertical block contains the categories up- and down-regulated. Colour intensity represents the percentage of genes belonging to each category and includes only GO terms for Biological Process. Clusters based on this percentage are displayed in different colours in the tree.

Figure 4. The *P. blakesleeanus madI* strains are heterokaryons of wild type and *madA* nuclei.

(A) Phenotype of wild type, *madA*, and *madI* strains L151 and L153, with illumination from the right (arrow).

(B) DNA sequencing chromatograms of the *madA* region in the two *madI* strains, and three progeny from each crossed to wild type UBC21 that represent the different outcomes in the progeny (Table S5).

(C) A segment of the sequence of MadA in different fungi with the conserved proline that is mutated in *madI* strains in bold.

(D) Graph showing the ratio of heterozygous/non-heterozygous SNPs in the genomes of 19 *mad* mutants.